2026-1466 IJOI
https://www.ijoi-online.org/

International Journal of Organizational Innovation

EARTHQUAKE MONITORING USING DEEP LEARNING WITH CNN–TRANSFORMER ARCHITECTURES

Yu-Chien Wu*[1]

*[1]Ph.D., Ph.D. Program for Infrastructure Planning and Engineering, Feng-Chia University, Taiwan. Email:gambler602238@gmail.com

Tse-Shan Hsu

President, Institute of Mitigation for Earthquake Shear Banding Disasters
Professor, Feng-Chia University, Taiwan.

Zong-Lin Wu*[2], Yan-Ming Wang*[1]

*[2]Assistant Professor, National Chin-Yi University of Technology, Taiwan.

Tsai-Fu Chuang, Da-Jie Lin

Associate Professor, Feng-Chia University, Taiwan.

Abstract

This study adopts the Chelungpu Fault and the disaster-affected Zhongxingxin Village in Taichung, Taiwan as case studies. We propose an earthquake monitoring framework based on a deep learning architecture. Strong-motion records are used to construct a near-fault training dataset and a nationwide testing dataset, enabling a comparative evaluation of four models, including a CNN–Transformer model. For near-fault earthquakes ranging from minor to moderate magnitudes, the results demonstrate that the CNN–Transformer model exhibits superior responsiveness and higher F1-scores in classification and identification tasks. Moreover, it shows the smallest performance degradation under cross-regional testing. The attention mechanism further enhances interpretability by effectively capturing the initial P-wave amplitudes and the characteristic energy frequency bands of S-waves. Consequently, under conditions where earthquake early warning is available, the proposed approach can contribute to earthquake disaster mitigation.

Keywords: earthquake, early warning, disaster mitigation, fault, CNN–transformer

Introduction

Taiwan is located on the Eurasian Plate and adjacent to the Philippine Sea Plate, resulting in frequent seismic activity. Although Taiwan is equipped with a comprehensive strong-motion monitoring network, severe earthquake damage occurred in Zhongxingsin Village during the 1999 Jiji earthquake, which was triggered by the rupture of the Chelungpu Fault. This observation indicates that (1) near-fault regions are exposed to higher levels of ground rupture and structural damage, and (2) variations in local site conditions among seismic stations hinder the generalization of AI-based models across different regions (University of Texas at Austin, 2024).

Therefore, this study adopts deep learning techniques (Mousavi et al., 2020; Hu et al., 2021; Lyu et al., 2025) to address these challenges by integrating spectral and time-domain features of seismic waves extracted using convolutional neural networks (CNNs) (Mousavi et al., 2020), long-term sequential dependencies learned through long short-term memory (LSTM) networks, and the sequence-weighting capability inherent in Transformer architectures. This integrated framework is employed to evaluate cross-regional discrepancies in model performance and to bridge the existing gap in AI-based earthquake monitoring models for near-field environments of active faults.

Study Area

The study area selected for this research is located in Zhongxing New Village, situated on the hanging wall of the Chelungpu Fault. The area is primarily composed of alluvial and colluvial deposits with weak cementation, which leads to significant amplification of seismic waves. During the 1999 Chi-Chi earthquake (Mw 7.6), pronounced localized surface ruptures were observed in Zhongxing New Village, where the maximum recorded seismic intensity reached Level 6.

Considering that the area is located on the hanging wall of the Chelungpu Fault, exhibits heightened sensitivity to ground motion, has a high density of residential buildings with elevated vulnerability to seismic hazards, and possesses abundant damage records from the 1999 Jiji earthquake, this region was selected as the field site for deep learning–based earthquake monitoring.

Seismic Data

This study utilizes continuous waveform records obtained from the Taiwan Strong Motion Instrumentation Program (TSMIP), operated by the Central Weather Bureau. These records are categorized according to different research objectives, as outlined below.

Training Data from Near-Fault Regions

The training data for near-fault regions are derived from seismic stations located within 0–15 km of the Chelungpu Fault. These data are intended to capture near-field seismic events generated by minor to moderate earthquakes, thereby enhancing the model's sensitivity to near-fault effects.

Test Data Across Taiwan

Test data from across the entire Taiwan region are used to evaluate the potential for deploying the models in different areas. The dataset includes seismic stations in northern, eastern, and southern Taiwan, encompassing a range of geological conditions, such as plains, basins, and mountainous regions. By adopting a dual-database approach, it is possible to compare the model's performance under "near-fault training" versus "cross-regional testing," thereby improving its practical applicability.

Data Preprocessing

Data Preprocessing Includes:

(a) Detrending and Noise Filtering

Removal of instrument drift and high-frequency noise.

(b) Sliding Window Segmentation

Division of continuous waveforms using a fixed time window.

(c) Normalization

Reduction of amplitude differences among stations.

(d) Multi-Channel Feature Construction;

(e) Acceleration Waveforms;

(f) Velocity Waveforms;

(g) Displacement Signals;

(h) STFT Spectral Features.

This set preserves both short-term transients in the seismic waves (e.g., those triggered by P-waves) and long-term energy variations (e.g., resulting from S-wave rupture and high-frequency amplification), enabling the model to distinguish differences in seismic source characteristics and earthquake magnitude.

Methodology and Model Architecture

A comparative study of CNN (Nakano and Sugiyama, 2022), LSTM, CNN-LSTM, and CNN-Transformer models will be performed. Here, CNN stands for Convolutional Neural Network, LSTM for Long Short-Term Memory, and CNN-LSTM for Convolutional Neural Network–Long Short-Term Memory. The sequence from input images to outputs is illustrated in the following flow chart.

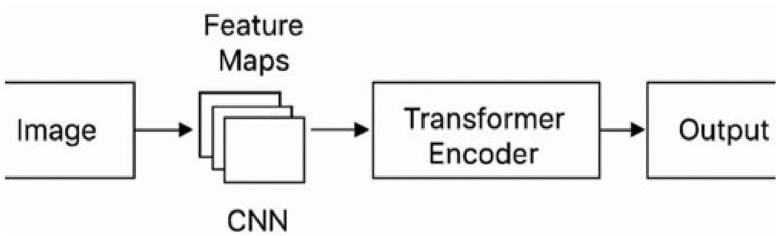

Figure 1. Flow chart of the sequence from input images to outputs.

Training Strategy and Loss Function

The model is trained using a supervised learning approach. The dataset is divided into training, validation, and test sets, and data augmentation techniques such as rotation, flipping, and scaling are applied to improve generalization. A mini-batch size of 32 is used, and the model parameters are optimized using the Adam optimizer with an initial learning rate of 0.001, which is decayed over epochs. Early stopping is applied based on validation loss to prevent overfitting, and regularization methods such as dropout are employed to enhance model robustness.

Seismic Wave Classification and Task Definition

In this study, continuous seismic waveforms are segmented into fixed time windows, and, according to practical requirements for strong-motion monitoring, events are classified into the following four categories:

Table 1. Classification of Seismic Events and Corresponding Definitions (N, S, M–L, D)

Seismic Events	Definition
N	Background noise
S	Mild near-field earthquakes
M-L	Moderate to strong near-field earthquakes
D	Distant strong earthquakes

The classification of seismic events and their corresponding definitions in Table 1 meets the requirements for rapid near-field identification, facilitating real-time earthquake assessment in densely populated areas (Wu, Khan, & Kwon, 2025) and enabling the timely detection of destructive seismic waves (Perol, Gharbi, & Denolle, 2018).

Model Comparison Framework

To compare the applicability of different deep learning architectures for near-fault monitoring, this study selects the following four models:

Table 2. The description, advantages, and limitations of each model (CNN, LSTM, CNN–LSTM, and CNN–Transformer) are summarized below..

Model	Description	Advantage	Limitation
CNN	Convolutional extraction of local time–frequency features	Proficient at capturing spectral pattern	Lacks long-range dependencies
LSTM	Long Short-Term Memory (LSTM) network	Capable of identifying time-series characteristics	Sensitive to noise and amplitude effects
CNN–LSTM	Local feature + sequence learning	Capable of handling both frequency-domain and time-domain processing	Unable to identify key segments
CNN–Transformer	Proposed in this study	With an attention mechanism to focus on critical bands	Computationally demanding

The core of this study is a CNN–Transformer model, which integrates the local feature extraction capabilities of CNNs with the sequential attention mechanism of transformers to improve near-fault detection performance and cross-domain generalization.

Hybrid CNN–Transformer architecture

The architecture proposed in this study mainly consists of the following components:

Convolutional Neural Network (CNN)

The CNN component is well-suited for extracting short-duration seismic waves and capturing spectral variations. It is particularly effective at identifying P-wave peaks while suppressing background noise.

Transformer Encoder

The Transformer encoder employs a multi-head self-attention mechanism to automatically assign weights to critical information within the seismic waveforms. For example, it can focus on the initial amplitudes of short P waves, enhance the energy rupture features of S waves, and ignore low-information noise segments.

Outputs

For multi-class classification, the model employs a Softmax function to output four categories of seismic events. Additionally, the model provides physically interpretable insights: the attention weights are primarily concentrated on the energy bands of P and S waves, which is consistent with established principles of earthquake engineering.

Definition and Interpretation of the F1 Score

True Positive (TP) refers to cases where the actual class is positive and the model correctly predicts a positive outcome. False Positive (FP) refers to cases where the actual class is negative but the model incorrectly predicts a positive outcome. False Negative (FN) refers to cases where the actual class is positive but the model incorrectly predicts a negative outcome. True Negative (TN) refers to cases where the actual class is negative and the model correctly predicts a negative outcome; it does not directly affect the F1 score.

Precision is defined as TP/(TP+FP), representing the proportion of true positive samples among all samples predicted as positive by the model. Recall is defined as TP/(TP+ FN), representing the proportion of actual positive samples that are correctly identified by the model.

The F1 score is calculated as

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}},$$

which is the harmonic mean of precision and recall. It balances the trade-off between these two metrics, preventing one from being excessively high while the other is too low.

Training Strategy

Loss Function: Weighted Cross-Entropy

A weighted cross-entropy loss is employed to mitigate classification bias caused by an excessive number of noisy samples in certain classes (N classes).

Optimizer: Adam with Adaptive Learning Rate

Regularization Methods:

1. Dropout: To prevent overfitting

Applied to prevent overfitting by randomly deactivating a fraction of neurons during training.

2. Data Augmentation: Noise augmentation, amplitude scaling

Techniques such as noise injection and amplitude scaling are employed to increase data diversity and improve model generalization.

3. Early Stopping: Overfitting

Training is terminated when the validation performance ceases to improve, preventing overfitting.

4. Window segmentation method

Consecutive overlapping sliding windows are employed to segment continuous seismic waveforms, simulating real-time seismic monitoring data.

Model design and training objectives

The model is designed to simultaneously optimize near-field sensitivity and cross-regional stability, ensuring it meets the requirements for real-world deployment.

Results

ClassificationRresults on the Near-Fault Database

Under the training and testing scenario using near-field data from the Chelungpu Fault, the CNN–Transformer model achieved the highest F1 scores across all four seismic wave categories. Notably, it showed significant improvements in recognizing mild near-field events (S class) as well as moderate to strong near-field events (M–L class). The average improvements compared with other models are as follows:

The model achieved approximately 3%–7% improvement over CNN–LSTM, 6%–12% improvement over CNN, and 8%–15% improvement over LSTM. These results demonstrate that the CNN–Transformer is more sensitive to near-field seismic waves characterized by rapid rupture and spectral variations, offering a significant advantage for the rapid identification of small to moderate near-field events in practical applications.

Wide Cross-Regional Applicability Across Seismic Stations in Taiwan

The wide cross-regional applicability of the model was evaluated to assess its potential deployment across different regions. Test data were collected from seismic stations in northern, eastern, and southern Taiwan. The results indicated the following:

- All models experienced performance degradation due to differences in site conditions.
- The CNN–Transformer exhibited the smallest performance degradation (less than 5%).
- The LSTM showed the largest degradation (approximately 18%).
- The CNN–LSTM exhibited intermediate degradation (approximately 8–10%).

Explanation

The LSTM relies heavily on amplitude trends, which reduces its stability under varying soil conditions. In contrast, the CNN–Transformer utilizes multi-head attention to extract physically meaningful features, resulting in superior cross-regional performance.

Attention Interpretability

Analysis of the model’s behavior using attention weights (Attention Maps) revealed that the CNN–Transformer automatically focuses on:

- S-wave energy concentrated within the 0.5–6 Hz frequency band.
- Exclusion of long-duration, low-information noise.

This interpretability is consistent with established engineering principles and facilitates adoption in disaster prevention systems, as it provides a transparent explanation for the model's decisions rather than relying on black-box predictions.

By analyzing the model's attention weights (Attention Maps), it was observed that the CNN–Transformer automatically focuses on:

- S-wave energy concentrated within the 0.5–6 Hz frequency band.
- The exclusion of long-duration, low-information noise.

This interpretability aligns with established engineering principles and supports adoption in disaster prevention systems, providing a transparent explanation for the model's decisions rather than relying on black-box predictions.

Real-Time Monitoring Performance

In this study, sliding windows were employed to simulate real-time seismic monitoring, and the inference computational efficiency is summarized as follows:

Table 3. The Results for Each Category of Average Inference Time and Decision Time

Category	Results
Average inference time	0.035～0.075 seconds
Decision time	Approximately 15–33 seconds before P-wave arrival

Significance: The model demonstrates real-time recognition capabilities, enabling rapid seismic assessment and supporting disaster prevention and mitigation in regions with densely distributed active faults.

Discussion

Advantages of CNN–Transformer in Near-Fault Monitoring

The CNN–Transformer model proposed in this study offers the following advantages for near-fault earthquake identification:

1. Cross-Regional Adaptability

The attention mechanism enables the model to automatically focus on seismically meaningful segments, reducing reliance on amplitude variations from a single station. This minimizes overfitting caused by differences in soil layers or terrain across Taiwan.

2. Capability to Identify Small-Scale Near-Field Seismic Waves

Mild near-field events are critical for early warning (Lyu, 2025) and situational awareness but are easily confused with noise. The CNN–Transformer demonstrates a superior ability to capture fine-grained seismic features, resulting in outstanding classification performance for S-class and M–L-class events.

3. Engineering Interpretability

Attention weights reveal that the model's decisions are primarily based on the short, high-frequency initial pulses of P waves and the energy-concentrated frequency bands of S waves. This behavior aligns with earthquake engineering principles regarding rupture processes and seismic wave propagation.

a. Cross-Regional Adaptability

The attention mechanism allows the model to automatically focus on seismically meaningful segments, reducing reliance on amplitude differences from a single station. This helps prevent overfitting caused by variations in soil layers or terrain across Taiwan.

b. Capability to Identify Small-Scale Near-Field Seismic Waves

Mild near-field events are critical for early warning (Lyu, 2025) and situational awareness but can be easily confused with noise. The CNN–Transformer demonstrates a superior ability to capture fine-grained seismic features, resulting in particularly strong classification performance for S-class and M–L-class events.

c. Engineering Interpretability

Attention weights reveal that the model's decisions are primarily based on the short, high-frequency initial pulses of P waves and the energy-concentrated frequency bands of S waves. This behavior is consistent with earthquake engineering principles regarding rupture processes and seismic wave propagation.

Limitations and Future Researches

Although this study demonstrates the feasibility of the proposed framework, several limitations can be addressed in future research:

1. Integration with Magnitude Regression

Currently, the framework performs only event classification. Incorporating magnitude prediction or peak ground acceleration (PGA) estimation could enhance the comprehensiveness of disaster mitigation applications.

2. Incorporation of Geotechnical Parameters

Near-fault monitoring remains influenced by geological and soil conditions. Including geotechnical parameters such as shear-wave velocity (Vs) and site effects could improve the model's accuracy for on-site, engineering-oriented deployment.

3. Challenges in Remote Deployment

Real-time monitoring is affected by communication latency, equipment reliability, and the logistics of on-site deployment. Automated systems require validation through collaboration with governmental seismic monitoring programs to

Although this study has demonstrated the feasibility of the proposed framework, the following limitations can be addressed in future research:

a. The classification categories have not yet been integrated with magnitude regression

Currently, only event classification has been performed; magnitude prediction or PGA (peak ground acceleration) estimation could be incorpo rated to enhance the comprehensiveness of disaster mitigation applications.

b. Geotechnical parameters (e.g., V_s, site effects) have not been incorporated

Near-fault monitoring is still affected by geological and soil conditions; incorporating geotechnical parameters could potentially improve the model's accuracy for on-site, engineering-oriented deployment.

c. Remote deployment in practice is still affected by communication latency

Real-time monitoring requires consideration of data transmission, equipment reliability, and on-site deployment. Automated systems still need to be validated in collaboration with governmental seismic monitoring programs.

Engineering Application and Deployment Potential

The results of this study indicate that deploying the CNN–Transformer deep learning model in near-fault, densely populated areas (such as Zhongxing New Village or urban development zones) could offer the following benefits:

1. Rapid seismic assessment

Provides real-time earthquake information to smart city monitoring platforms.

2. Decision support for institutions

Can serve as a reference for building safety evaluations and disaster management center judgments.

3. Early warning potential

Capable of identification 15–33 seconds before P-wave arrival, serving as an auxiliary tool for earthquake early warning.

Conclusions

This study, using Zhongxing New Village and the Chelungpu Fault as examples, proposes a deep learning model combining CNN and Transformer (**3) for seismic monitoring and rapid identification in active fault zones. The main conclusions are as follows:

1. CNN–Transformer demonstrates advantages in near-fault identification

CNN–Transformer demonstrates advantages in near-fault identification, with both classification sensitivity and F1-scores outperforming CNN, LSTM, and CNN–LSTM. Its improvement is particularly notable for the recognition of mild to moderate near-field events, supporting more accurate seismic assessment in densely populated areas.

2. Demonstrates cross-regional applicability

In tests conducted across seismic stations throughout Taiwan, the model exhibited a performance degradation of less than 5%, indicating that it is not significantly affected by variations in soil layers, geological conditions, or amplitude differences, and thus is feasible for deployment in diverse regions.

3. Physical interpretability of attention weights

The model concentrates on the high-frequency signals at the onset of P-waves and the energy-concentrated frequency bands of S-waves, consistent with earthquake rupture mechanics and seismic wave propagation characteristics, thereby enhancing the applicability of deep learning in engineering applications.

4. It demonstrates potential for real-time monitoring and early warning applications

Through the use of sliding windows and rapid inference, the model can generate earthquake classification judgments within approximately 15–33 seconds after the arrival of P-waves, demonstrating practical value for integration with smart city disaster prevention and near-fault monitoring systems.

Overall, the proposed CNN–Transformer in this study serves as a concrete tool for near-fault earthquake monitoring and rapid identification, with potential for practical engineering applications and system integration.

References

Hu, Y., Zhang, Q., Zhao, W., & Wang, H., “TransQuake: A transformer-based deep learning approach for seismic P-wave detection,” *Earthquake Research Advances,* Volume 1, Issue 2, https://doi.org/10.1016/j.eqrea.2021.100004, 2021

Kasburg, V., Müller, J., Eulenfeld, T., Breuer, A., & Kukowski, N., “Cross-Regional Seismic Event Discrimination via CNNs: ensemble averaging and fine-tuning methodsm,” *Bulletin of the Seismological Society of*

America, Volume 113, Issue 2, pp. 842–856, 2023

Lyu, D., Nakata, R., Ren, P., Mahoney, M. W., Nakata, N., & Erichson, N. B, "*Rapid wavefield forecasting for earthquake early warning via deep sequence to sequence learning)*," *Nature Communications, 16*, 10622. https://doi.org/10.1038/s41467-025-65435-2*,* 2025

Mousavi, S. M., William L. Ellsworth, Weiqiang Zhu, Lindsay Y. Chuang, Gregory C. Beroza, "Earthquake Transformer — An Attentive Deep-Learning Model for Simultaneous Earthquake Detection and Phase Picking,*" Nature Communications*, *11:3952, https://doi.org/10.1038/s41467-020-17591-w,* 2020

Nakano, M. and Sugiyama, D., "Discriminating seismic events using 1D and 2D CNNs: applications to volcanic and tectonic datasets," *Earth, Planets and Space,* Volume 74, Article 134. https://doi.org/10.1186/s40623-022-01696-1, 2022

Perol, T., Gharbi, M., and Denolle, M., "Convolutional neural network for earthquake detection and location," *Science Advances,* Volume 4, Article 2, https://doi.org/10.1126/sciadv.1700 578, 2018

University of Texas at Austin, "Jackson School researchers employ AI to predict earthquakes in China," Retrieved from https://global.utexas.edu/news/jackson-school-researchers-employ-ai-predict-earthquakes-china, 2024

Wei Li, W., Chakraborty, M., Köhler, J., Quinteros-Cartaya, C., Rümpker, G., & Srivastava, N., "Earthquake monitoring using deep learning with a case study of the Kahramanmaras Turkey earthquake aftershock sequence," *Solid Earth,* Volume 15, pp. 197–213, 2024

Wu, A., Khan, I., and Kwon, Y.-W.' "A transformer-based real-time earthquake detection framework in heterogeneous environments," Voulme 15, Article 1, 8422. https://doi.org/10.1038/s41598-024-82087-2, 2025